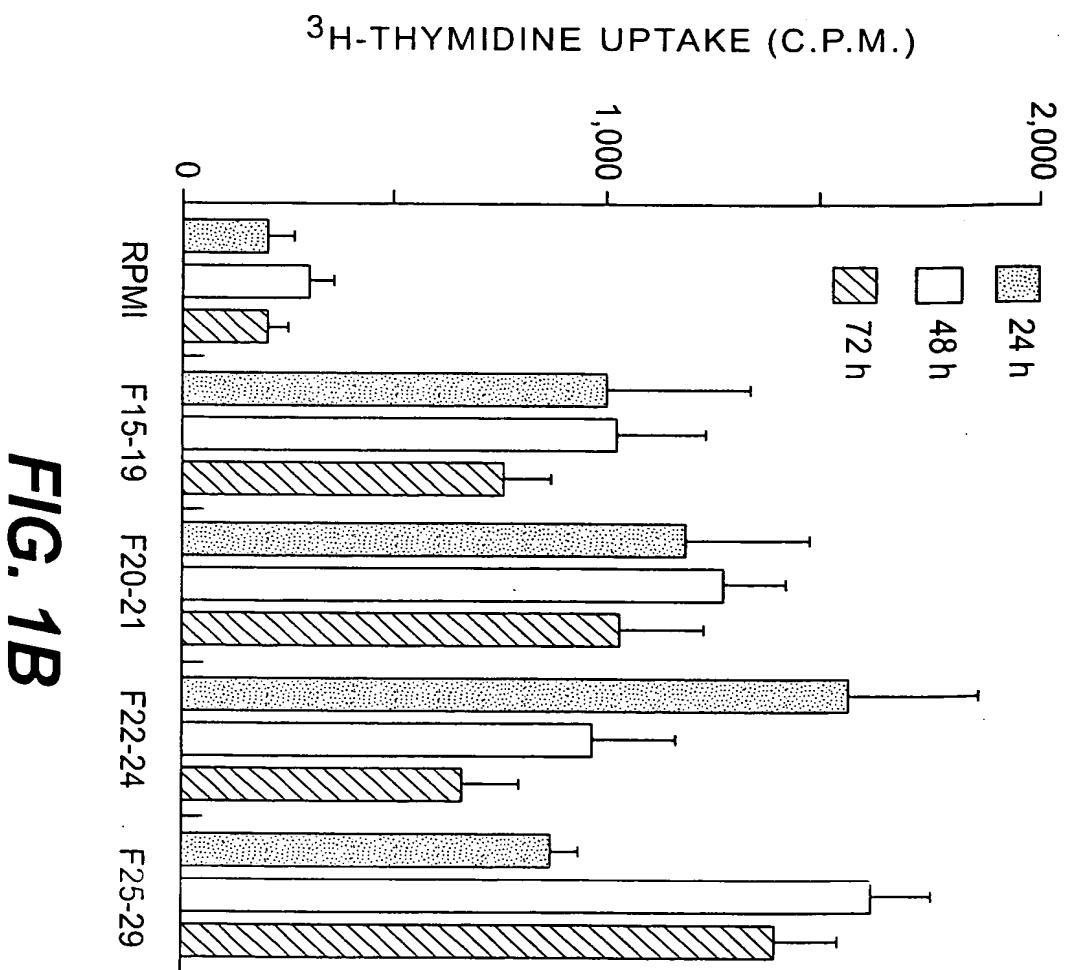
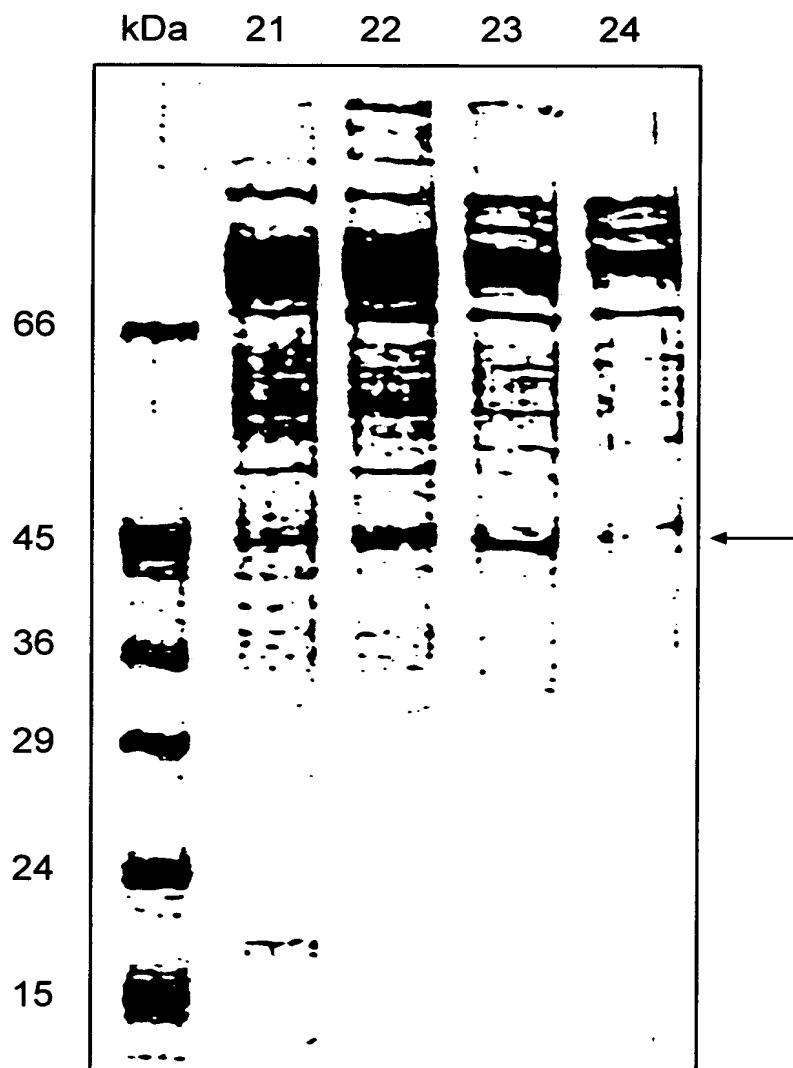


**FIG. 1A**



**FIG. 1B**

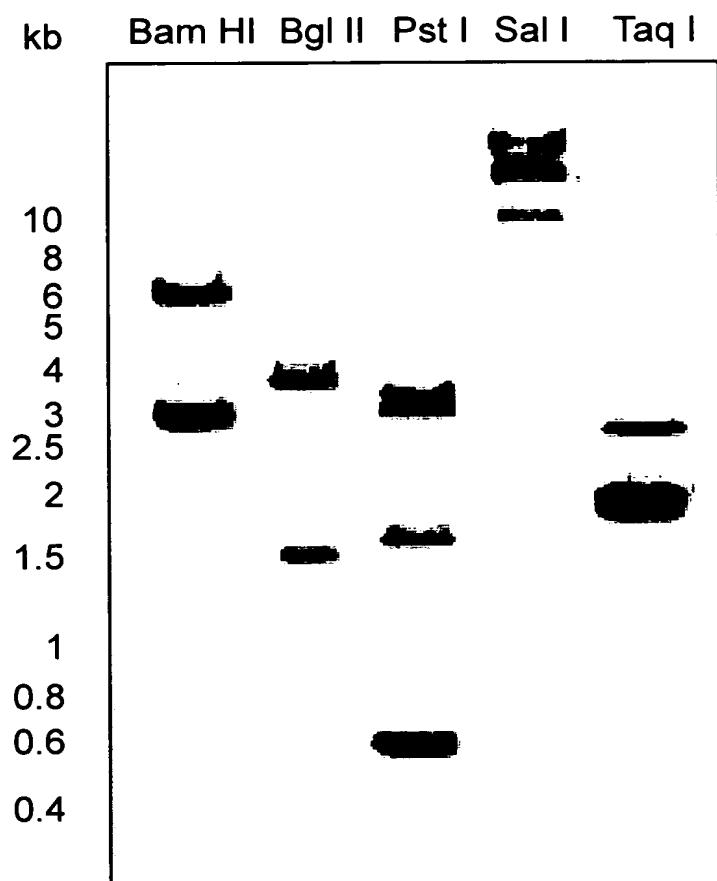
HPLC FRACTIONS



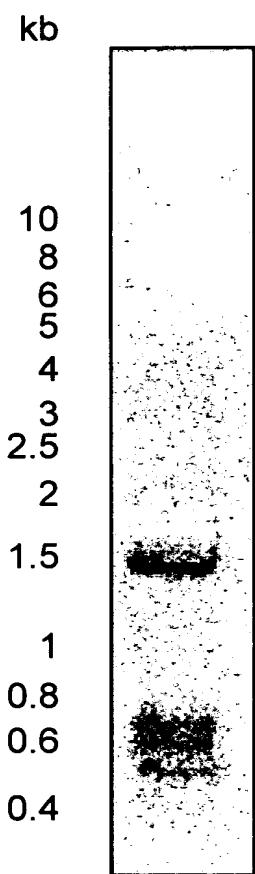
**FIG. 1C**

TC	MRKSVCVKQFFFSAFPFFFFCVFPLISRTQKQYKIIKGEEKKKNQRANRREHQKREIMRFKKS	75
CS	-----MKFSKG	6
Pa	-----MQR	3
TC	FTCIDMHTGEAARIWTSGLPHIPGSNMAEKAYLQENMDYLRRGIMLEPRGHDDMFGAFLDPFEEGADLGWVF	150
CS	IHAIDSHTMGEPTRIVWGGIPQINGETMADKKYLEDNLVDYVRTALMHEPRGHNDMFGSIITSSNNKEADFGIIF	81
Pa	IRIDSHTGGEPTRLVIGGFPLGQGDMERRRLIGERHDAMRAACILEPRGSDVLVGALLCAPDPEACAGVIF	78
TC	MDTGGYLNMCGHNSIAAVTAAWETGIVSVPAKATNVVULDTPAGLVRGTAHLOSGTESEVNASIINVPSFLYQ	225
CS	MDGGGYLNMCGHGSIGAATVAVEATGMVEMVEPVTNIN--MEAPAGLIKAKVMVEN--EVKVEVSITNVPSFLYQ	151
Pa	FNNSGYLGMCGHGTIGLVALASLAHLGRIGPGV----HRIETPVGEWEATLH-----EDGSVSVRNVPAYRYR	140
TC	ODVVVVLPKPYGEVRVDIAFGGNFAIVPAEOLGIDISVONLSRLOEAGELLRTEINRSVKVOHPOLPHINTVDC	300
CS	EDAKLEVPSLNKTTFDISFGGFFFAIHAKELGVKVETSQDVLLKKLGIEIRDLINEKIKVQHPELEHIKTVDL	226
Pa	RQSVSEVPGI--GRVSGDIAMGGGNWFFLVAGH--GQLAGDNLDALTAYTWAVOQALDD---QDIRGEDGAIIDH	208
TC	VEIYGPPTNPEANYKNVVFIGNMROADR [SPCGT] GTSAKMATLYAKGQLRIGETFYYESILGSLFQGRV--LGEE	371
CS	VEIYDEPSNPEATYKNVWTFGQQVDR [SPCGT] GTSAKLATLYKKGHLKIDEKFVYESITGTMFKGRV--LEET	297
Pa	IELFAD--DPHADSRNFVLCPGKAYDR [SPCGT] GTSAKLACLAADGKLIPGQWPWRQASVIGSQFEGRYEWLDQ	279
TC	RIPGVKVPVTKDAEEMLVVTAETGKAFIMGFTNTMLFDPTDPFKNGFTLKQ* 423	
CS	KVGEFD-----AIIPEITGGAYITGFNFHVVIDPEDPLKYGFTV*-- 335	
Pa	PGGPPIVPTIRGRAHVSAEATLLADDDPFANGIRR*----- 314	

**FIG. 2**

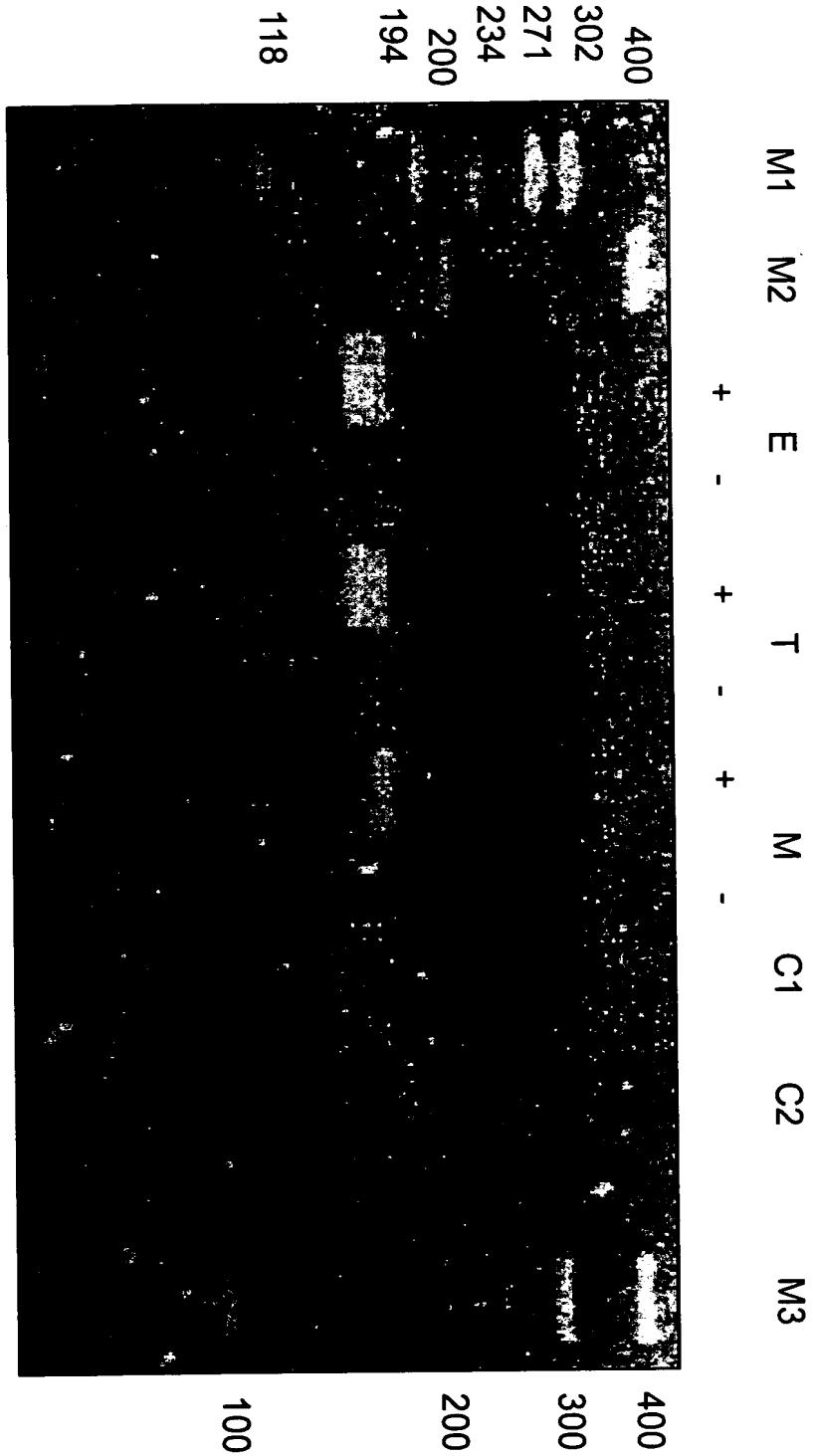


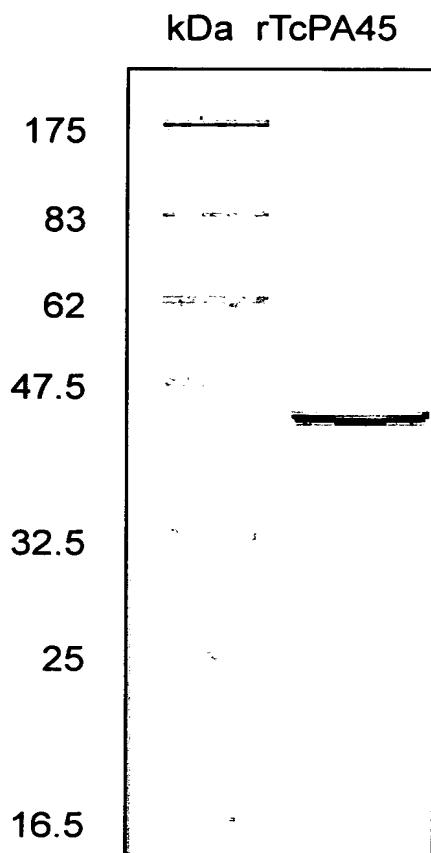
**FIG. 3A**



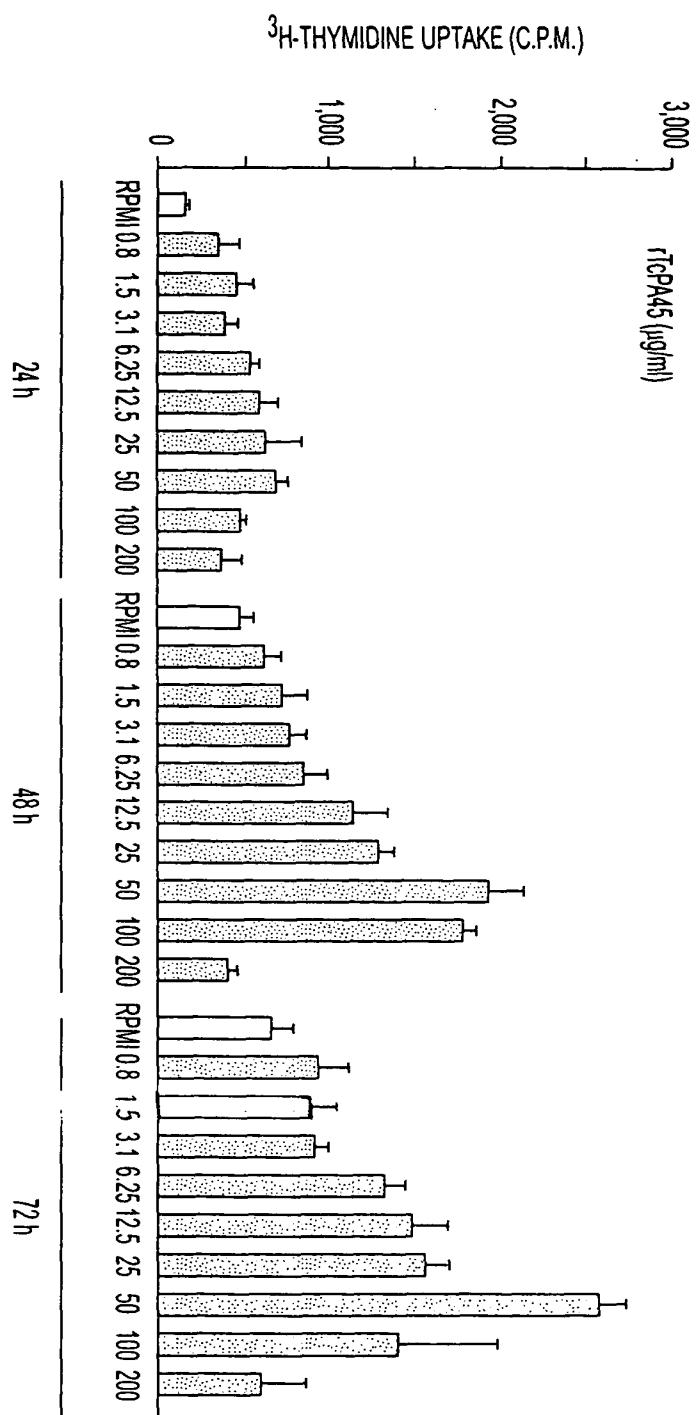
***FIG. 3B***

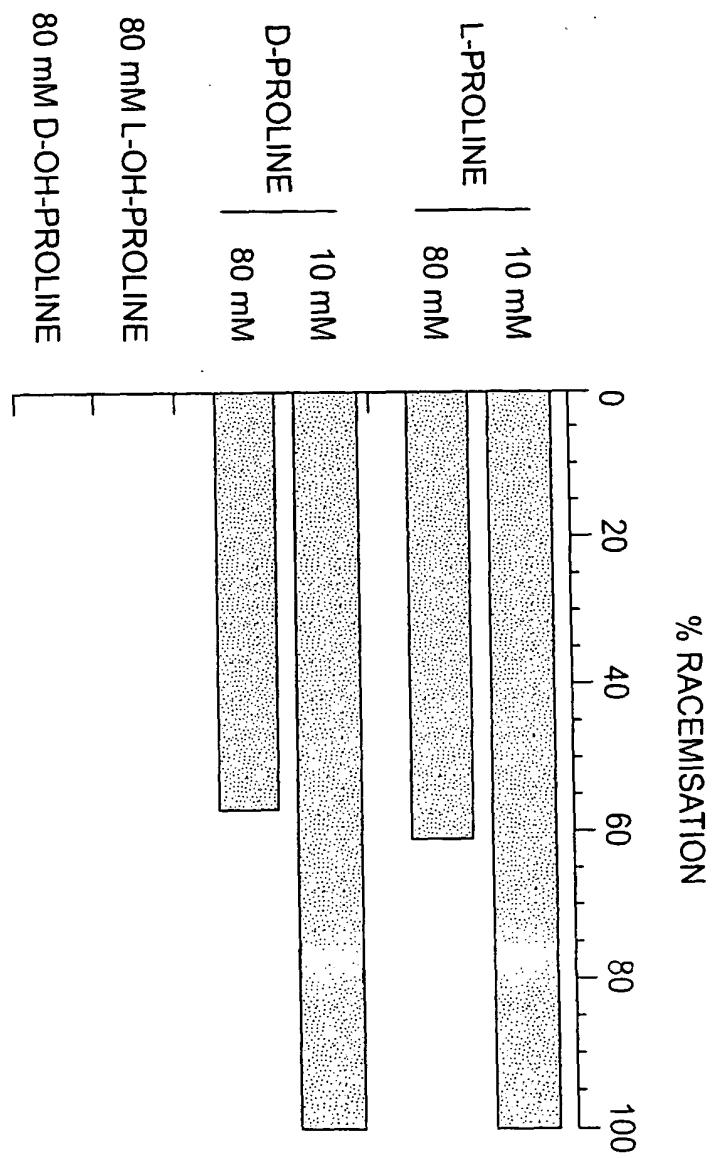
**FIG. 3C**



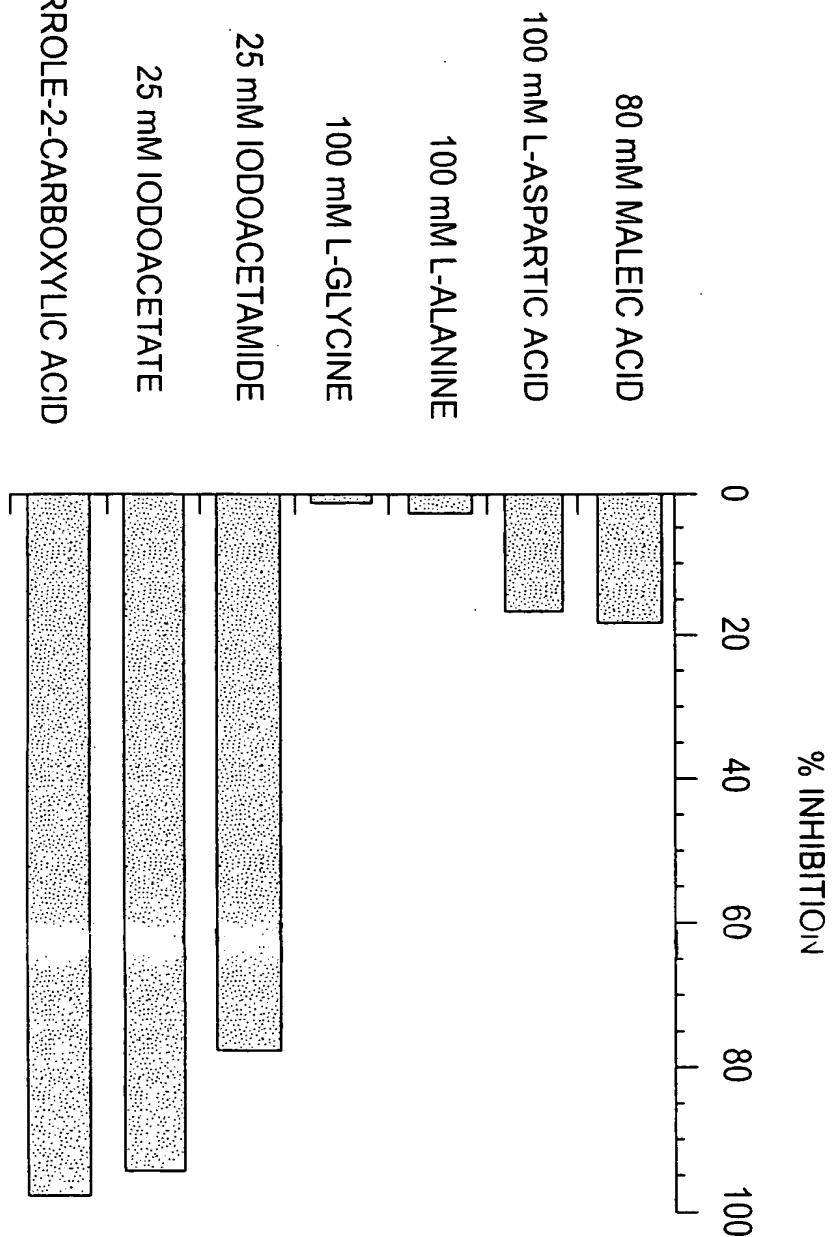


**FIG. 4A**

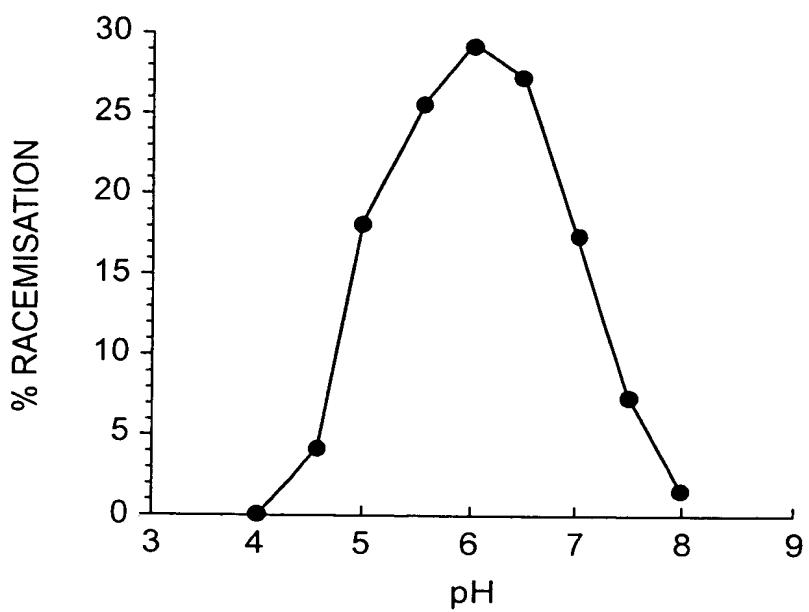




**FIG. 4C**



**FIG. 4D**



**FIG. 4E**

POLYPYRIMIDINE RICH REGION

SPLICE LEADER  
 ACCEPTOR SITES

SIGNAL

CCTTTTCTTTAAACAAAAATTCCGGGGAAATATGGAACAGGGTATAGGCTAAAGTGTCTGTC<sub>CAAACAAA</sub>TTTT  
 TTTCCGCCTTCCCATTTTTTTGTGTGTTCCCTGATCTCGAACAGGGCAGGAAGCTCTGTTGCCAAAT 90  
 F S A F P F F F C V F P L I S R T G Q E K L L F D Q K Y 12  
 AAAATTAAGGCGAGAAAAAGAAAAAAATCACGAGCAACAGGAGAACACCACAAAGGAAAT 180  
 K I I K G E K K K N Q R A N R R E H Q Q K R E I M R F 42  
 AGGAATCATTCACATGCATCGACATGCATACGGAAAGGTGAAGCAGCACGGATTGTGACGAGTGGTTGCCACACATCCAGGTCGAAT 270  
 K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 72  
 ATGGCGGAGAACGATACTGCAGGAACATGGATTATTGAGGCCACGGTGGCATAATGCTGGAACCACGGT 360  
 M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 102  
 GGAGCCTTATTGACCTATGAAGAACGGCGCTGACTGGCATGGTATTCATGGATACGGGGCTATTAAATATGTGTGGACAT 430  
 G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 132  
 AACTCAATTGCAGGTTACGGGGCACTTGAAACGGGAATGTGAGCGTGGGGCAACAAATGTTCCGGTTGTCCTGGACACA 520  
 N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T 162  
 CCTGGGGGTGGTGCGGTACGGCACACCTCAGAGTGGTACTGAGAGTGGTCAAATGCAGTATTTAATGTACCCCTCATT 610  
 P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 700  
 TTGTATCAGCAGGATTGGTGGTGTGCCAAAGCCCTATGGTGAAGTACGGGTGATATTGCAATTTTTCGCCATT 790  
 L Y Q D V V V V L P K P Y G E V R V D I A F G G N F F A I 222  
 GTTCCCGCGAGCAGTGGAAATGTATCTCCGTCAAACCTCTCCAGGCTGCAGGAGCAGGAACTTCTGCGTACTGAAAT 252  
 V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 880  
 282

**FIG. 5A**

CGCAGTGTGAAGTTAACGCACCCCTCAGCTCCCCATATAACACTGTGGACTGTGTGAGATATCGGTCGCCAACGACCCGGAGCA 970  
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 312  
AACTACAAGAACGTTGATATTGGCAATGCCAGGGATGCCCTCCATGTGGCACAGGCCAACATGGCAACACTTAT 1060  
N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 342  
GCCAAAGGCCAGCTGGCATCGGAGAGACTTTGTGTACGGAGAGCATACTCGGCTCACTCTCCAGGGAGGGTACTTGGGAGGGAGA 1150  
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 372  
ATACCGGGGCTGAAGGTGCCGGTGACCAAAGATGCCGGAGGAAGGGATGCTCGTGTAA CGGCAGAAATTACTGAAAGGTTTATCATG 1240  
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 402  
GGTTCAACACCATGCTGTTGACCCAACGGATCCGTTAAGAACGGATTCCACATTAAAGCACTGAGTAGATCTGGTAGAGCACAGAACTATT 1330  
G F N T M L F D P T D P F K N G F T L K Q 423  
GGGAACACGTCGCAACAGGTGCTACGTGAAGGGTATGAATGAACTGGTTTTTATTATTATTATTAGTAGTCATT 1420

ATATTAATTTTTTGTGTTGGGTTCAACGGTACCGGCTGGGAGCAGGGAGCGGATAGGGGGACAATTGGCTTTAT 1510

TTCATTTCATCTCCTACCCACCCCTGGTCCACCGGTGCCGGGGCTGTGGGTGGAGGACTCTAAATCCGCACCTCGG 1600

AGGAATAACATATTCAATTCTCATCTTGAATCAAAGGCAT

1651

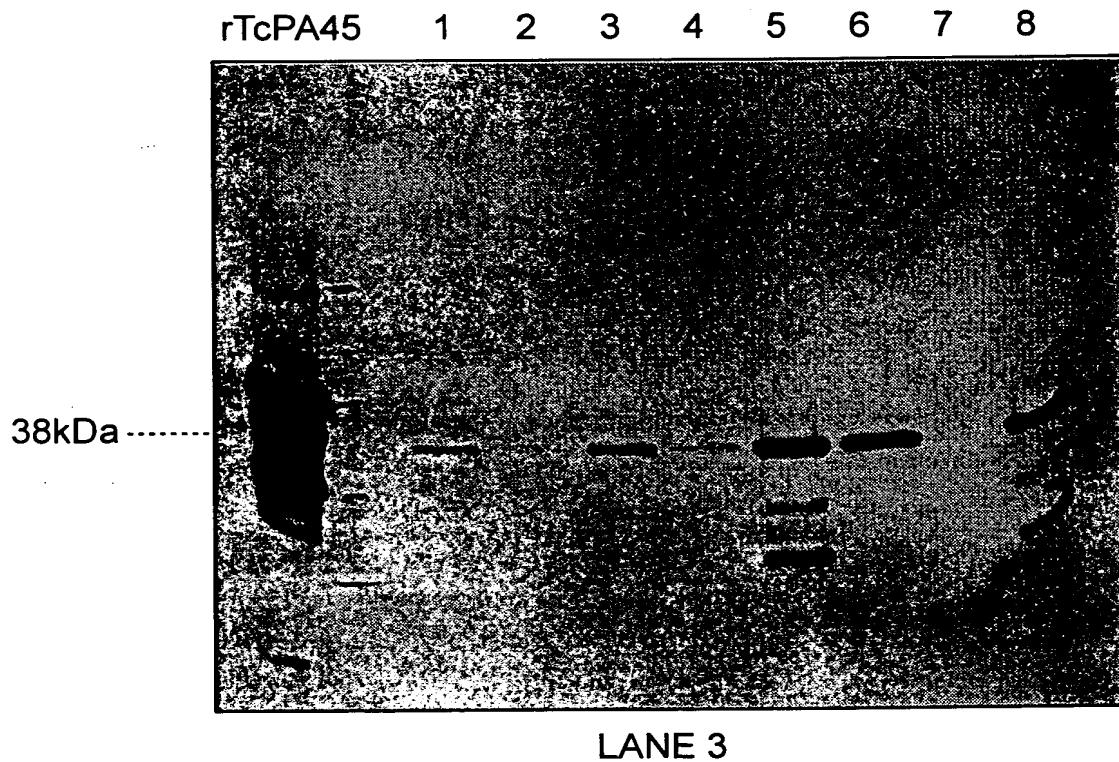
POLYADENYLATION SITE

OBS: UNDERLINED THE SEQUENCED PEPTIDES USED TO DEDUCE DEGENERATED PRIMERS FOR CLONING

NUCLEOTIDE SEQUENCE AND PEPTIDE SEQUENCE TCPA45

**FIG. 5B**

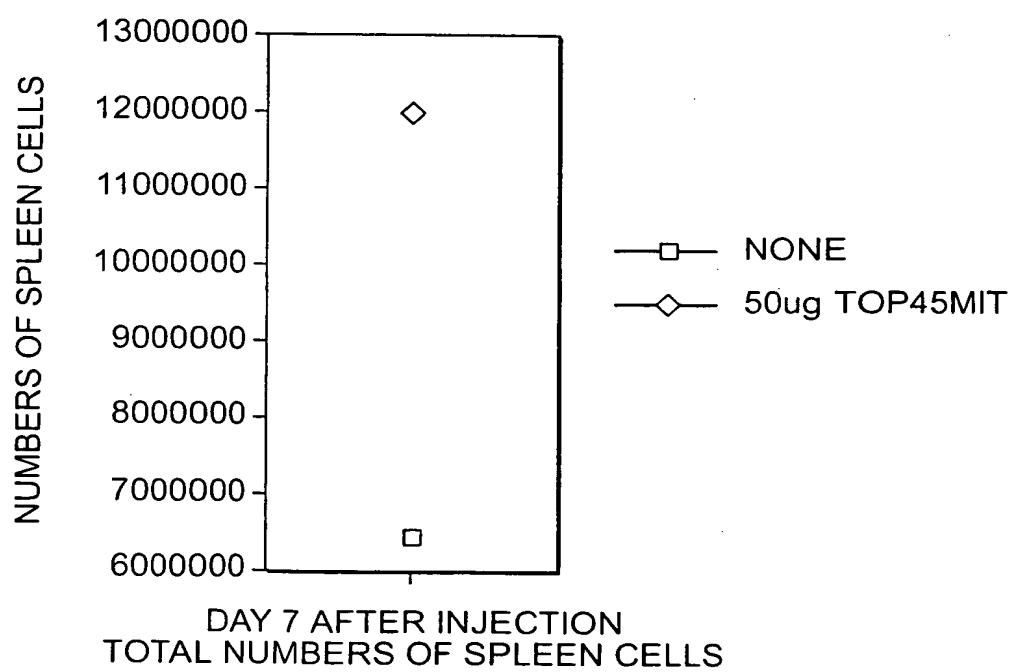
WESTERN BLOTH



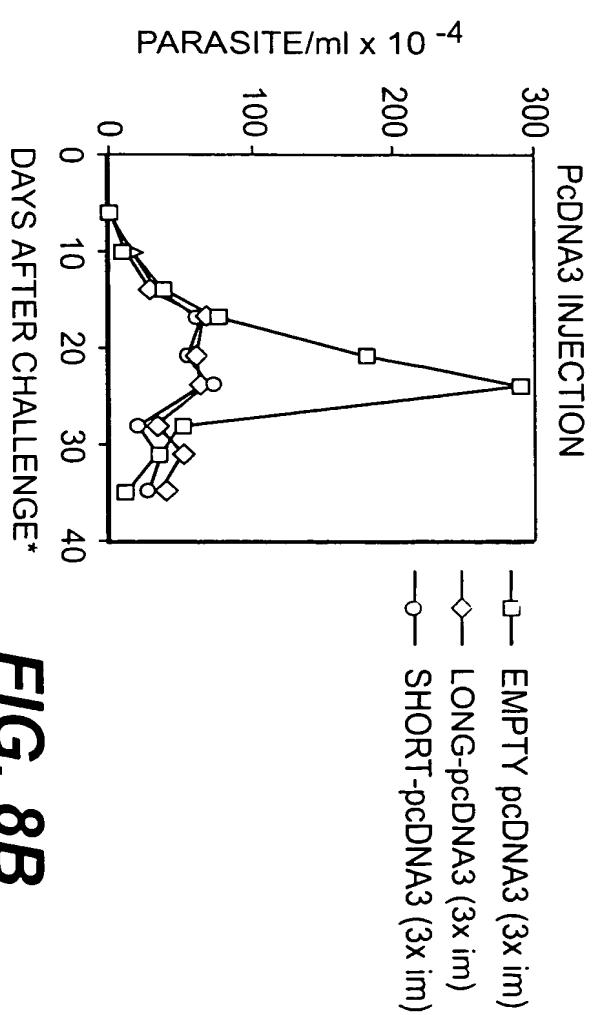
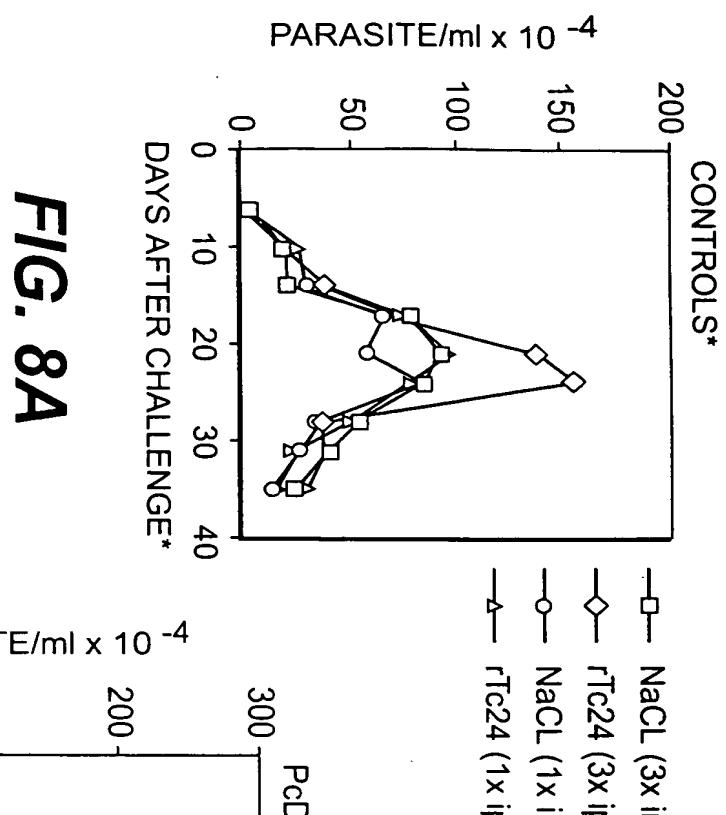
SOLUBLE FRACTION OF EPIMASTIGOTES EXTRACT (CYTOSOLIC)  
REVEALED WITH ANTIBODY DIRECTED TO rTcPA45

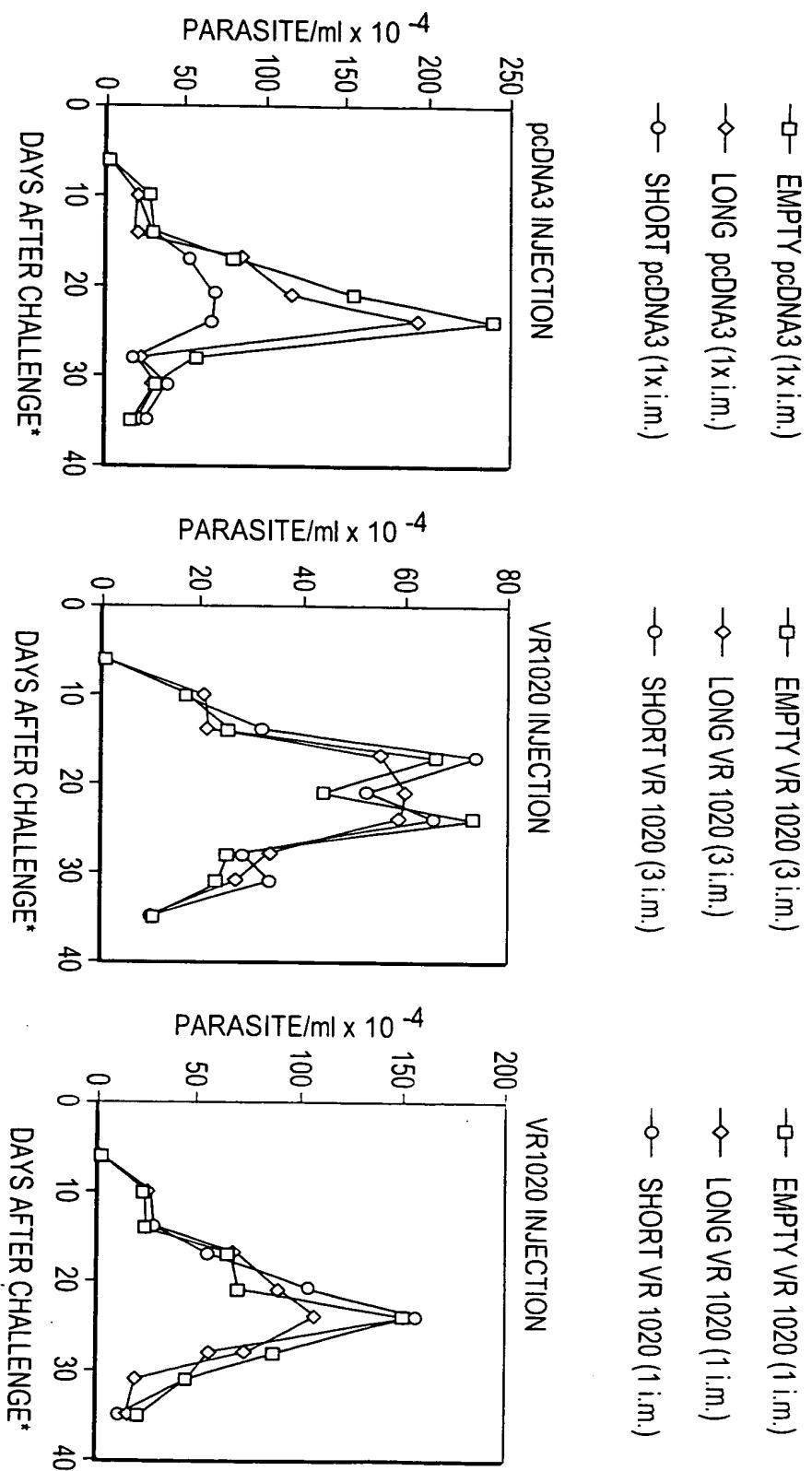
----- DEMONSTRATES THE EXISTANCE OF AN INTRACYTOPLASMIC  
FORM OF TcPA45 IN THE PARASITE

**FIG. 6**



**FIG. 7**

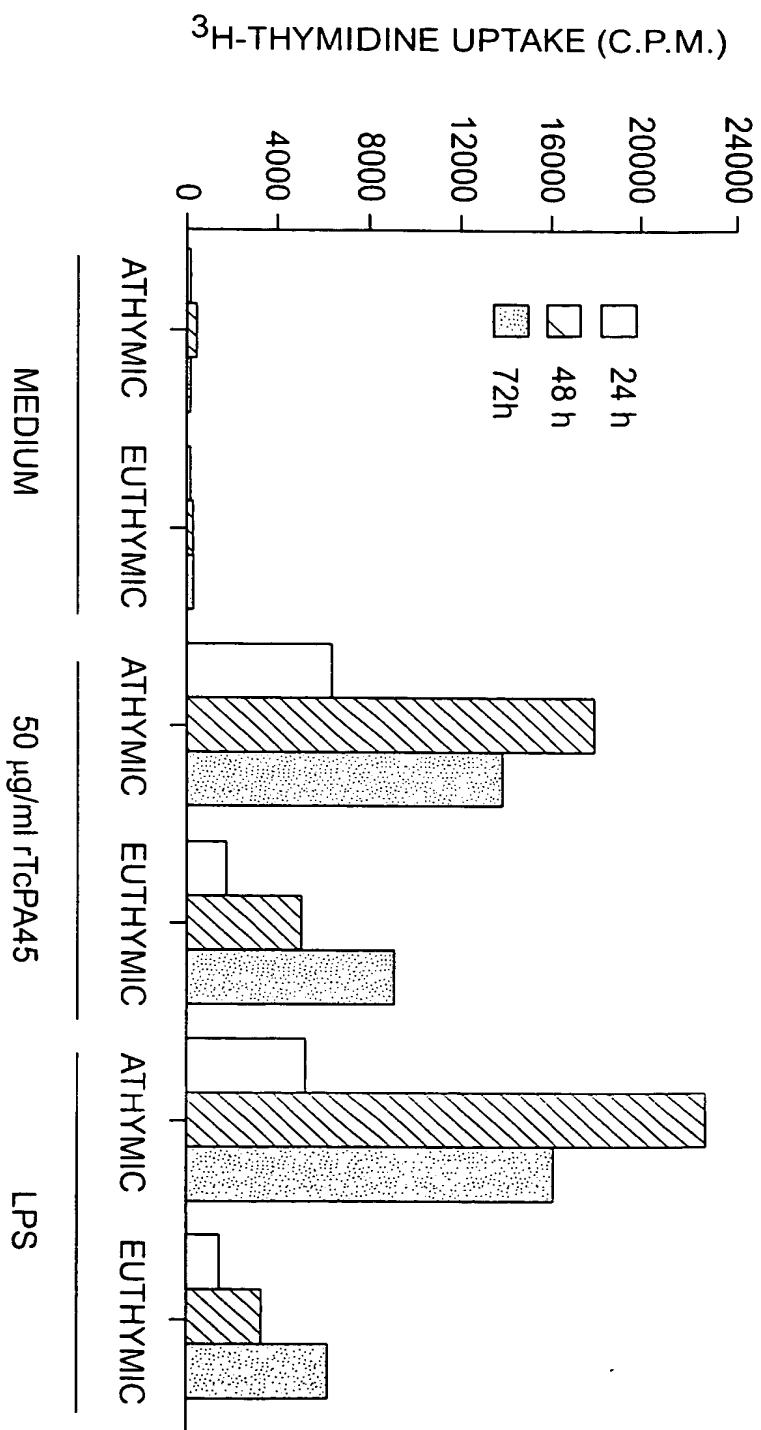




**FIG. 8C**

**FIG. 8D**

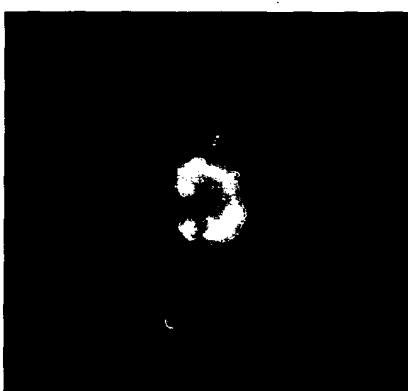
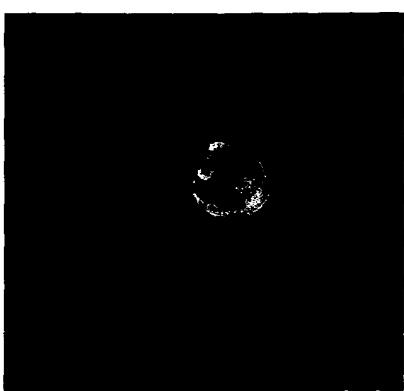
**FIG. 8E**



**FIG. 9**

Alexa-F (ab')2

Chronic serum

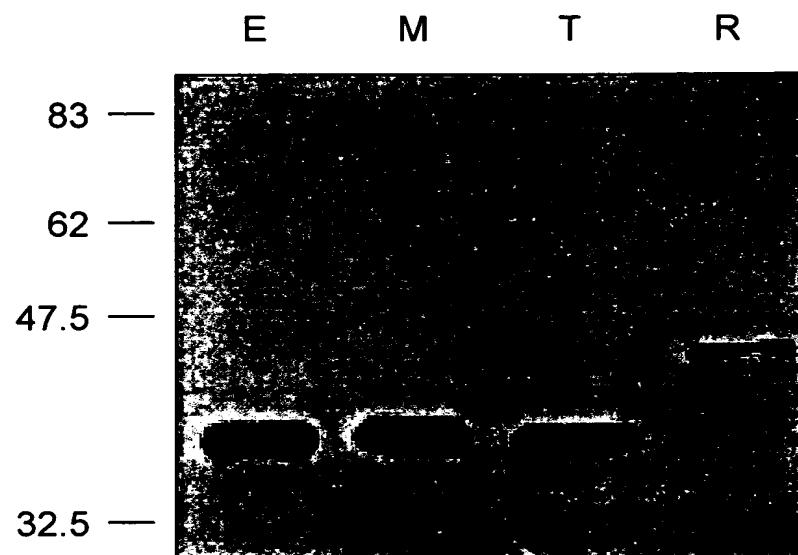


EPIMASTIGOTE

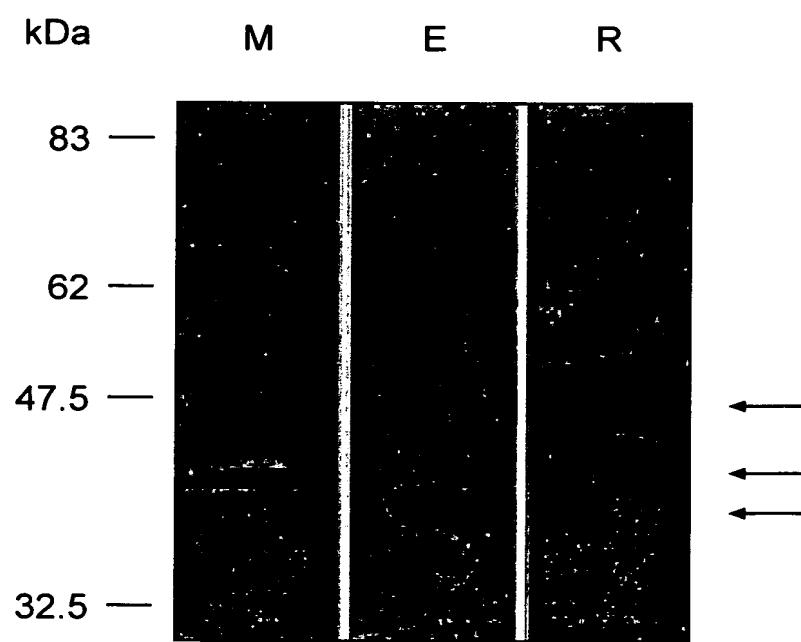
METACYCLIC

TRYPOMASTIGOTE

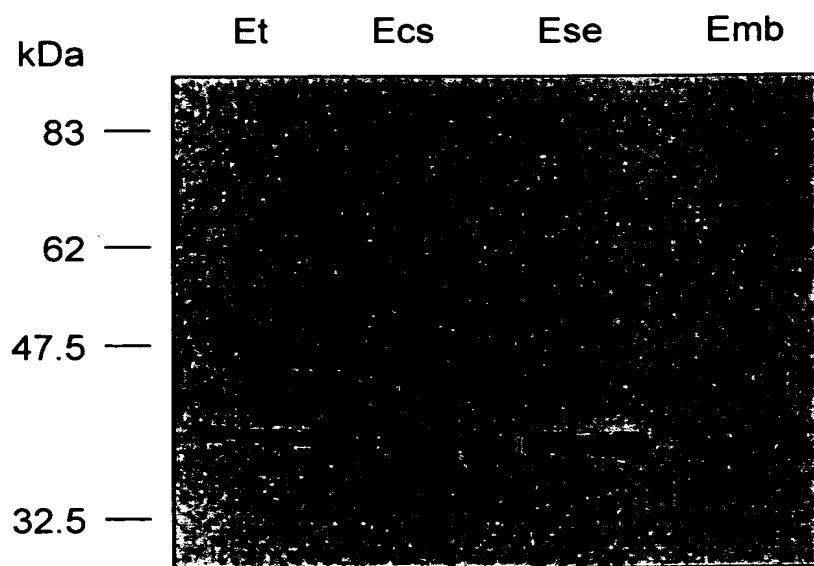
**FIG. 10A**



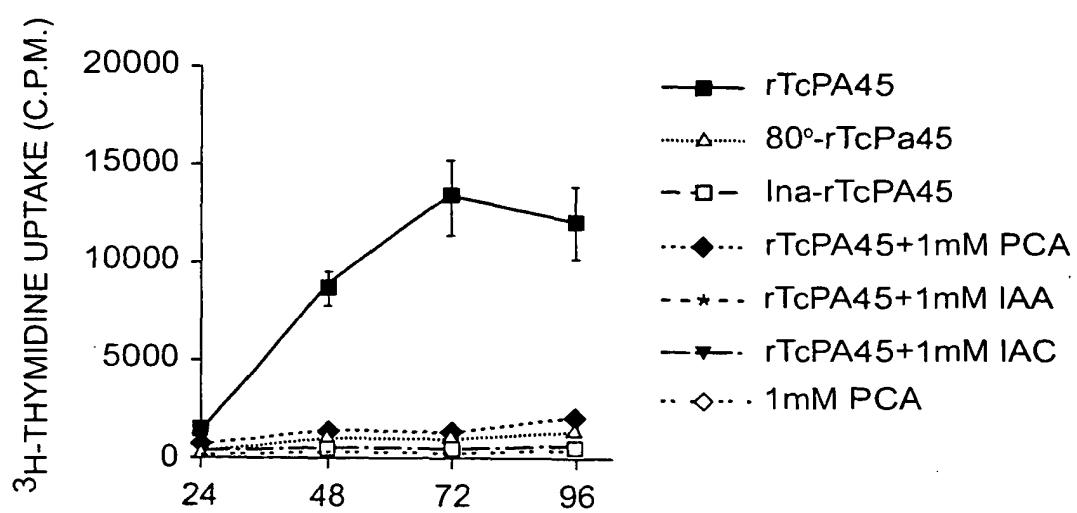
**FIG. 10B**



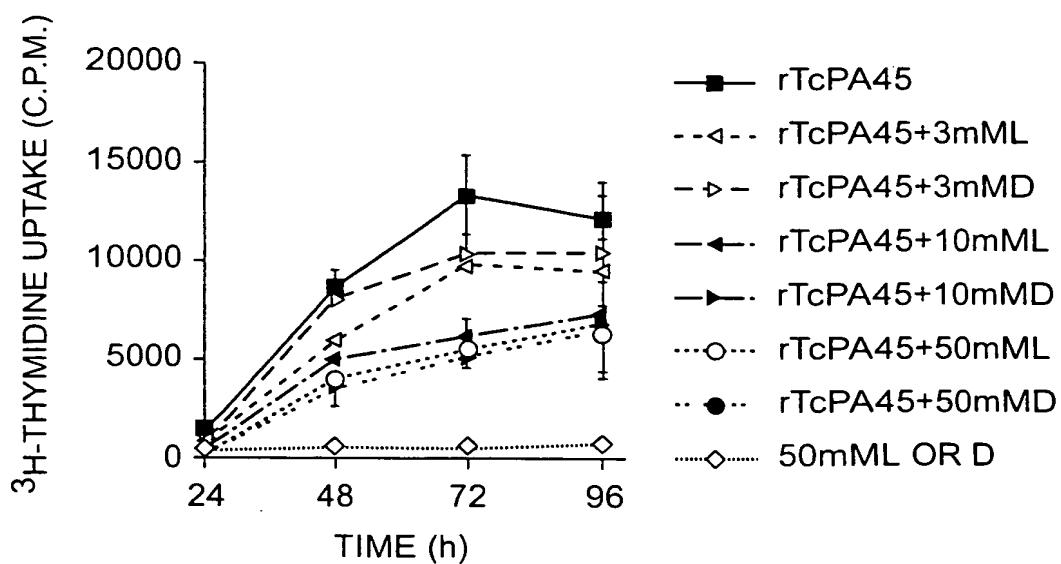
**FIG. 10C**



**FIG. 10D**

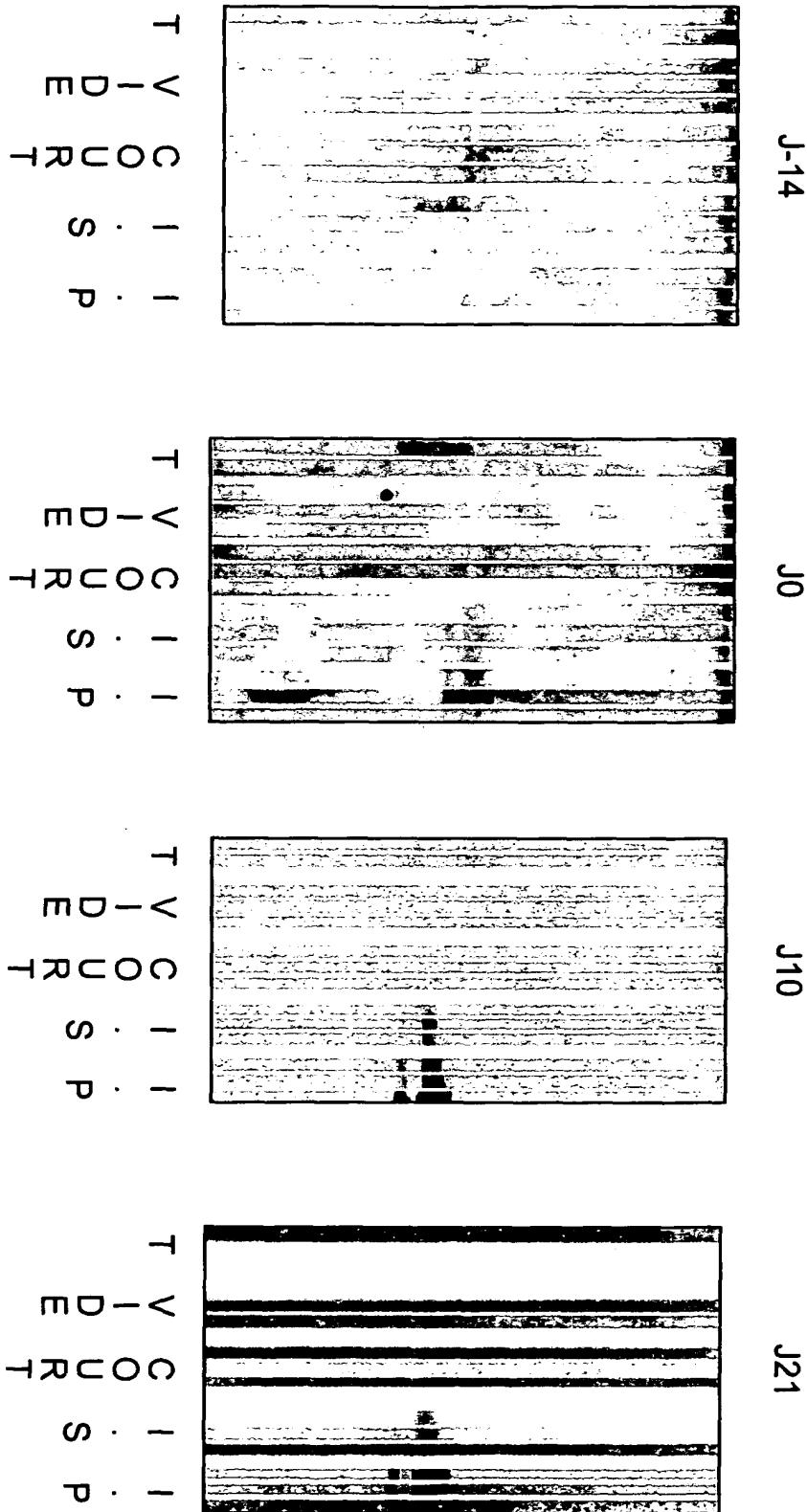


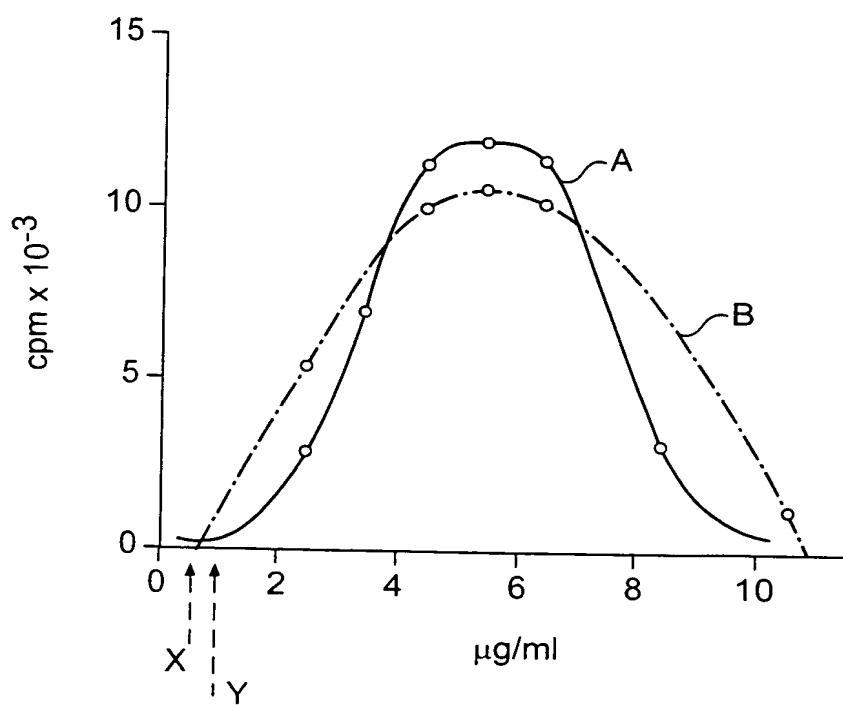
**FIG. 11A**



**FIG. 11B**

**FIG. 12**

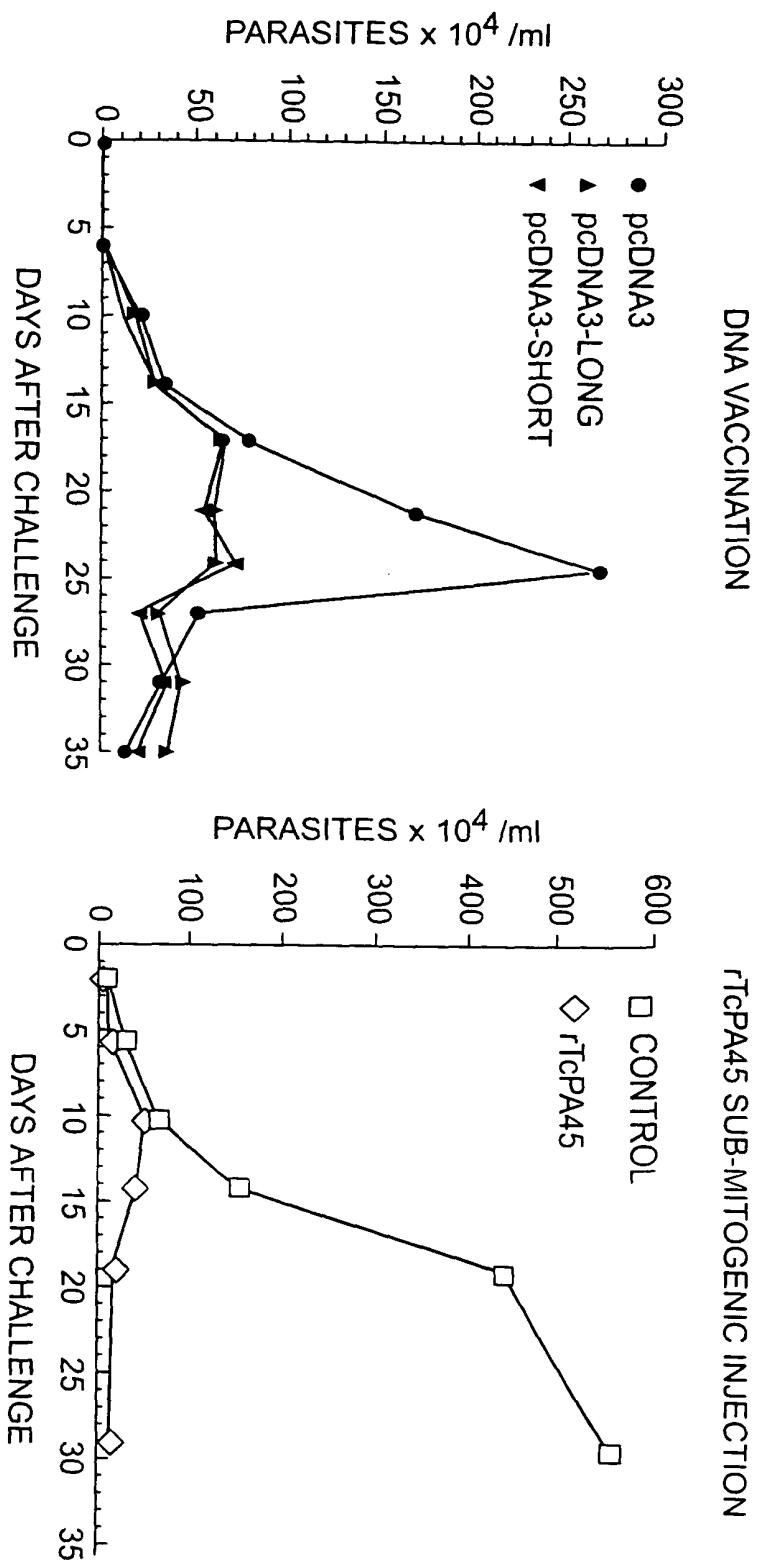




**FIG. 13**

**Replacement Figure 14A and 14B  
Divisional Application of 09/725,945**

Inventors: Minoprio et al.  
Attorney Docket No. 03495-0200-02000



*FIG. 14A*

SEQ ID NO:2

TC RTGQEKL~~LFDQKYK~~LIKGEKKKNQRANRREHQKREIMRFKKS 75  
TC FTCIDMHTEGEAARIVTSGLPHIPGSNMAEK~~KAYLQENMDYLLRGIMLEPRGHDDMFGAFLFDPIEGADLG~~MVF 150  
TC MDTGGYLNMCGHNSIAAVTAAVETGIVSVPAKATNVPVLDTPAGLVRGTAHLSQSGTESEVSNASIINVPSFLYQ 225  
TC QDVVVVLPKPYGEVRVDIAGGNFAIVPAEQIGIDISVQNL~~SRIQEAGE~~ELLIRTEINRSVKVQH~~POLPHINTVDC~~ 300  
TC VEIYGPPTNPEANYKNVWIFGNQADR SPCGT GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE 371  
TC R~~PGVKV~~PVTKDAE~~GMLV~~VTAET~~GKA~~IMGFTMLFDPTDFKNGFTLKQ\* 423

**FIG. 15**

SEQ ID NO:4

Tc	MRFKKS	75
Tc	FTCIDMHTEGEAA <u>RIVTSGLPHIPGSNMAEK</u> KAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPIEEGADLGMVF	150
Tc	MDTGGYLNMC <del>G</del> HNSIAAVTAAVETGIVSVPAK <u>ATNWVWLDTPAGLVRG</u> TAHLQSGTESEVSNASIINVPSFLYQ	225
Tc	QDVVVWVLPKPYGEV <u>RVDIAFGGN</u> FAIVPAEQLIGIDI SVQNLISRLQEAECELLRTEINRSVKVOHPQLPHINTVDC	300
Tc	VEIYGPPTNPEANYK <u>NWVIFGNRQADR</u> SPCGT GTSAKMATLYAKQLRIGETFVYESILGSLFQGRV--LGEE	371
Tc	RIPGVKVVPVTKDAEEGMLVVTAEITGKAFIMGFNTMLFDPTDPFKNGFTLKQ*	423

**FIG. 16**

SEQ ID NO:7

POLYPYRIMIDINE RICH REGION



**SPLICE LEADER  
ACCEPTOR SITES**

**SIGNAL PEPTIDE**

CCTTTTCTTTAAAACAAAAATTCCGGGGGAATGGAAC <u>AGGCTATATGCC</u>	<u>TAAAGTGCTGTCCC</u>	<u>AAACAAAATT</u>	TTT	90
<u>TTTCCGCCTCCCCATT</u>	<u>TTTGTGTTCCCTGATCTCGAACAGGCAGGAAACCTCTGTTGACCAAAAT</u>	M	R K S V C P K O K F F	12
F S A F P F F F C V F P L I S R T G Q E K L L F D Q K Y				42
AAAATTAAAGGGCGAGAAAAAGAAAAAAATCACCGAACAGGAGAACACCAACAAAAAGGAAATT <u>TGCGATT</u>	K I I K G E K K E K K N Q R A N R R E K Q Q R R E I M R F			270
AAGAAATCATCACATGCACATGCACATGGAAAGGTGAAGCAGCACGGAT <u>TGTGACCGAGTGGTTGCCACACATCCAGGTCGAAT</u>	K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N			360
ATGGCGGAGAGAACGATACTGCAGGAAACATGGATTATTGAGGCGTGCATAATGCTGGAAACCACGGTGCATGATGATA <u>TGTT</u>	M A E K K A Y L Q E N M D Y L R R G I M L E P R G K D D M F			430
GGAGCCTTTATTGACCCATTGGAGAACGGCGCTGACTGGCATGGTA <u>TTCATGGATACCGGTGGCTATTAAATATGTTGACAT</u>	G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H			520
AACTCAATTGCAGCGGTACGGCGCA <u>GGTGAACGGGAATGGTGGCGTGGCGAACAAATGTCGGTTGTCCTGGACACA</u>	N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T			610
CCTGGGGGT <u>TGGTGCACACCTCAGAGTGGTACTGAGAGTGGTCAAATGGAGTATATCAATGTACCCATT</u>	P A G L V R G T A R L Q S G T E S E V S N A S I I N V P S F			700
<u>TTGTATCAGCAGGATGGGGTGTGTC</u> <u>AAAGCCCTATGGTGAAGTACGGGTGATATTGCATTGGAGGCAATTTCGCCATT</u>	L Y Q Q D V V V V L P K P Y G E V R V D I A F G G N F F A I			790
				222
				252

**FIG. 17A**



SEQ ID NO.8

<u>ATGGCTAAAGTGTCTCCCACAAACAAATT</u>	90
M R K S V C P K O K F F	12
<u>TTTCCGCCTTCCCATTTCCTTGTGTTCCCTGATCTCGAACAGGCAGGAAAGCTCTGTGACCAAAAT</u>	180
F S A F P F F F C V F P L I S R T G Q E K L L F D Q K Y	42
<u>AAATTATAAGGCCGAGAAAAAGAAAAAGAAAAATCACCGACAAACAGGAGAACACCAACAAAAGGAATTATGCGATT</u>	270
K I I K G E K K E K K N Q R A N R R E H Q Q K R E I M R F	72
<u>AAGAAATCATCACATGCATCGACATGCCATACCGAAGGTGAAGCACCACGGATTGTGACGACTGTTGCCACACATCCAGGTTCGAAT</u>	360
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N	102
<u>ATGCCGGAGAGAACGATACCTCAGGAAACATGGATTATTGAGGCCATATACTGGAAACCACGTGGCATGATGATATGTT</u>	430
M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F	132
<u>GGAGCCTTTTATTGACCTATTGAAAGAGGCCTGACTTGGCATGGTATCATGGATACCGGGCTATTAAATATGTTGACAT</u>	520
G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H	162
<u>AACTCAATTGAGCGGTACGGCGCAGTGAAACGGAATTGTGAGCGTGGCGAGGCAACAAATGTCGGTTGCTGGACACA</u>	610
N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T	192
<u>CCTGGGGGTGGTGGCGGTACGGCACACCTCAGAGTGGTACTGAGAGTGGTCAAATGGAGTATTACAATGTACCCATT</u>	700
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F	222
<u>TTGTATCAGCAGGATGGTGGTGTGTTGCAATTGGAGTATGGGAATTGGCAATTTCGGCATT</u>	790
L Y Q Q D V V V L P K P Y G E V R V D I A F G G N F F A I	252
<u>GTTCCCGCGGAGCAGTGGAAATTGATATCTCCGTCACAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTCTCGCTACTGAATCAAT</u>	880
V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N	282
<u>CGCAGTGTGAAGGTTCACGCACCCCTAGCTGCCCATATTAACTGTGGACTGTGAGATATACGGTCCCAACGAAACCGGGAGGCA</u>	970
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A	312

**FIG. 18A**

AACTACAAGAACGTTGTATTTGCCAATGCCAGGGATCGCTCCATGGGCACGGCACCAGGCCAAGATGCCAACACTTAT 1060  
N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 342  
GCCAAAGGCCAGCTTCGCATGGAGACTTGTGTACCGAGAGCATACTCGGCTCACTCTCCAGGCAGGGTACTGGGGAGGCGA 1150  
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 372  
ATACCGGGGTGAAGGTGCCGGTACCAAGATGCCGAGGAAGGGATGCTCGTGTAAACGGCAGAAATTACTGGAAGGCTTATCATG 1240  
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 402  
GGTTCAACACCATGCTGTTGACCAAACGGATCCGTTAAGAACGGATTCACATTAAGCACTGAGTAGATCTGTTAGAGCACAGAACTATT 1330  
G F N T M L F D P T D P F K N G F T L K Q \* 423  
GGGGAACACACGTGCGAACAGGTGCTACGTGAAGGGTATTGAATGCTTTTATTATTATTATTATTAGTCATT 1420

ATTATTAATTTTTTTTGTTTGGGTTAACGGTACCCGTTGGGACCGGAGCGATAGGGGGACAATTTCGTTT 1510

TTTCATTTCATCTCCTACCCACCCCTTGGTCCACCGTCCGGGGCTCTGGGTGAGGACTCTAACTCCGACCTCGG 1600

AGGAATAACATATTCAATTCAATCTGGAAATCAAAGGCAT

1651

POLYADENILATION SITE

Obs: UNDERLINED THE SEQUENCED PEPTIDES USED TO DEDUCE DEGENERATED PRIMERS  
FOR CLONING

NUCLEOTIDE SEQUENCE AND PEPTIDE SEQUENCE TcPA45

**FIG. 18B**

SEQ ID NO:9

CGAACAGGGCAGGAAAGCTCTGTTGACCAAAAT 270  
R T G Q E K L I F D Q K Y 72  
AAAATTAAAGGGCGAGAAAAAGAAAAAAATCAACGGCAAACAGGAGAACACCAACAAAAAGGAAATTTGCCGATT 360  
K I I K G E R K E K K N Q R A N R R E H Q Q R R E I M R F 102  
AAGGAATCATCACATGCATCGACATGCATAACGGAAAGGTGAAGCCAGCAGGATGTGACCAGTGGTTGCCACACATTCAGGTTCGAAT 430  
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 132  
ATGGCGGAGAGAACATACTGCAGGAAACATGGATTATTGAGGCCATATGCTGGAACCACGGGGTCATGATGATATGTT 520  
M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 162  
GGAGCCTTTATTGACCTATGAGAAGGGCTGACTTGGCATGGTATCATGGATAACGGGGCTATTAAATATGTTGGACAT 610  
G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 192  
AACTCAATGCAGCGGTACGGCGCAGTTGAAACGGGAATGTGAGCGTGGCGAACAGCAACAAATGTCCGGTGTCCGGACACA 700  
N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T 222  
CCTGGGGGTGGTGGCGGTACGGCACACCTCAGAGTGGTACTTGAGAGTGGTCAAATGGAGTATTATCAATGTACCCATT 790  
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 252  
TTGTATCAGCAGGATGGTGGTGTGTTGCCAAAGCCCTATGGTAAGTACGGGTGATATTGCAATTGGAGGCAATTTCGCCATT 880  
L Y Q Q D V V V L P K P Y G E V R V D I A F G G N F F A I 282  
GTTCCCGGGAGCAGTGGAAATTGATATCTCCGTCACAAACCTCCAGGCTGCAGGAGGCAGAGACTCTGGTACTGAATCAAT 970  
V P A E Q I D I S V Q N L S R L Q E A G E L L R T E I N 312  
CGCAGTGTGAAGGTTCAAGCACCCCTCAGCTGCCCATATTAACACTGTGGACTGTGAGATATCGGTCCGCCAACGAAACCCGGAGGCA 1060  
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 342  
AACTACAAGAACGTTGTGATATTGGCAATGCCAGGGCATGCCCTCCATGTGGCACAGGCCAACAGATGCCAACACTTAT 1150  
N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 372

**FIG. 19A**

GCCAAAGGCCAGCTCGCATCGGAGAGACTTTGTGTACGGAGGCATACTCGGCTCACTCCAGGGCAGGGTACTGGGAGGGCGA 12400  
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 402  
 ATACCGGGGTGAAGGTGCCGGTGACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAGGCCATTATCATG 13300  
 I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 423  
 GGTTCACACCATGCTGTTGACCCAACGGATCCGTTAAGAACGGATTACATTAAGCAGTAGATCTGGTAGAGCACAGAACTATT 14200  
 G F N T M L F D P T D P F K N G F T L K Q \*  
 GGGAACACGTGCGAACAGGTGCTACGTGAAGGGTATGAATCGTTTTTATTTTATTTTATAGTGCAATT 15100  
 ATTATAAATTTTTTGTGTTGGGTTCAACGGTACCGTACCCAAACCCCTGGTCCACCGGTGGGACGGAGCCATAGGCCGGACAATTGCTTTAT 16000  
 TTTCATTTCATCTTCTACCCAAACCCCTGGTCCACCGGTGGGACGGAGCCATAGGCCGGACAATTGCTTTAT 1651

POLYADENYLATION SITE

**Obs: UNDERLINED THE SEQUENCED PEPTIDES USED TO DEDUCE DEGENERATED PRIMERS  
FOR CLONING**

NUCLEOTIDE SEQUENCE AND PEPTIDE SEQUENCE TcPA45

**FIG. 19B**

SEQ ID NO:10

SIGNAL PEPTIDE

ATGCGTAAAGTGTCTGTCCAAACAAATTTTT

TTTTCCGCCTCCCCATTTTTTTTTGTTGTTCCCTTGATCTCT

NUCLEOTIDE SEQUENCE OF SIGNAL SEQUENCE TcPA45

**FIG. 20**

SEQ ID NO:11

		ATGCCATT	360
	M	R	F
K	K	S	F
T	C	I	D
M	H	T	E
G	E	A	A
R	I	V	T
S	G	L	P
H	I	P	G
I	V	T	S
G	L	P	H
N	S	N	S
132	430	132	430
M	A	E	K
K	A	Y	L
Q	E	N	M
M	D	Y	L
R	R	G	I
L	E	P	R
E	P	R	G
H	D	D	M
D	M	F	F
162	520	162	520
GGAGCCCTTTTATTGACCCATTGAAGAACGGCTGACTTGGCATGGTATTCAATGGATACCGGTGGCTATTAAATAATGTCGTGACAT	610	GGAGCCCTTTTATTGACCCATTGAAGAACGGCTGACTTGGCATGGTATTCAATGGATACCGGTGGCTATTAAATAATGTCGTGACAT	610
G	A	F	L
F	D	P	I
I	E	E	G
G	A	D	L
A	D	L	G
M	V	F	M
V	F	M	D
T	G	G	Y
G	H	H	L
N	M	C	N
M	C	G	G
C	G	H	H
192	700	192	700
AACTCAATTGCAGCGGTACGGCGCACACCTTCAGAGTGGTACTGAGACTGAGGTGTCAAATGGAGTATTATCAATGTCCTGGACACA	700	AACTCAATTGCAGCGGTACGGCGCACACCTTCAGAGTGGTACTGAGACTGAGGTGTCAAATGGAGTATTATCAATGTCCTGGACACA	700
N	S	I	A
A	A	V	T
A	A	V	E
E	T	G	I
S	V	P	A
V	P	A	K
A	T	N	V
T	N	V	P
V	V	L	D
L	D	T	T
222	790	222	790
CCTGGGGTTGGTGCACGGCACACCTTCAGAGTGGTACTGAGACTGAGGTGTCAAATGGAGTATTATCAATGTCCTGGACACA	790	CCTGGGGTTGGTGCACGGCACACCTTCAGAGTGGTACTGAGACTGAGGTGTCAAATGGAGTATTATCAATGTCCTGGACACA	790
P	A	G	L
G	L	V	R
V	R	G	T
G	T	A	H
L	Q	I	Q
I	Q	S	G
S	G	T	E
T	E	S	E
E	S	N	A
S	N	A	S
A	S	I	I
I	I	N	V
N	V	P	S
V	P	S	F
252	880	252	880
TGTATCAGCAGGATGTTGGTGTGCCAAGCCATTGGTAAGTACGGGTGATATTGCCATTGGCAATTTCGCCATT	880	TGTATCAGCAGGATGTTGGTGTGCCAAGCCATTGGTAAGTACGGGTGATATTGCCATTGGCAATTTCGCCATT	880
L	Y	Q	Q
Q	D	V	V
V	V	V	L
L	P	K	P
K	P	Y	G
Y	G	E	V
G	E	V	R
E	V	R	V
V	D	I	A
D	I	A	F
I	A	F	G
A	F	G	N
F	A	I	I
282	970	282	970
GTCCCCGGAGCAGTTGGATATCTCCGTCAAACCTCCAGGCCAGGGAGGAACTTCTGGTACTGAATCAAT	970	GTCCCCGGAGCAGTTGGATATCTCCGTCAAACCTCCAGGCCAGGGAGGAACTTCTGGTACTGAATCAAT	970
V	P	A	E
E	Q	L	G
Q	L	G	I
L	D	I	S
D	I	S	V
I	S	V	Q
S	Q	N	L
Q	N	L	S
N	L	S	R
L	Q	E	A
Q	E	A	G
E	A	G	E
A	G	E	L
G	E	L	L
E	L	L	R
L	R	T	E
R	T	E	I
T	V	D	C
V	D	C	V
D	C	V	E
C	V	E	I
V	E	I	Y
E	I	Y	G
I	Y	G	P
Y	G	P	P
G	P	P	T
P	T	N	P
T	N	P	E
N	P	E	A
P	E	A	342
342	1150	342	1150
AACTACAAGAACGGTGTGATATTGGCAATGCCAGGGATGCCATGGGACAGGCCAACAGGCCAACATGGCAACACCTTAT	1150	AACTACAAGAACGGTGTGATATTGGCAATGCCAGGGATGCCATGGGACAGGCCAACAGGCCAACATGGCAACACCTTAT	1150
N	Y	K	N
Y	K	N	V
K	N	V	V
N	V	V	I
V	I	F	G
I	F	G	N
F	G	N	R
G	N	R	Q
N	R	Q	A
R	Q	A	D
Q	A	D	R
A	D	R	S
D	R	S	P
R	S	P	C
S	P	C	G
P	C	G	T
C	G	T	G
G	T	G	T
T	G	T	S
G	T	S	A
T	S	A	K
S	A	K	M
A	K	M	A
K	M	A	T
M	A	T	L
A	T	L	Y
402	1240	402	1240

**FIG. 21A**

ATACCCGGGGTGAAGGGGCCGTCACCAAAGATGCCGAGGAAGGGATGCTCGTTGTAACGGCAGAAATTACTGGAAAGGCTTTATCATG 1330  
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 423  
GCTTCAACACCATGCTGTTGACCAACGGATCCGTTAACGAGTAGATCTGGTAGAGCACAGAACTATT 1420  
G F N T M L F D P T D P F K N G F T L K Q \*  
GGGAAACACGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAAATCGTTTTTATTATTATTATTAGTGCATT 1510  
ATTATAAATTTCATTTGGGTTCAACGGTACCGCTGGGAGCAGGAAGCGATAGGCCGACAATTTCGCTTTAT 1600  
TTTCATTTCATCTCCTACCCAACCCCTTGTTCCACCGGTGGGGGGTCTTGTTGGAGGACTCAAATCCGCACCTCGG 1651  
AGGAATAACATATTCAATTTCATATCTTGAATCAAAGGCA

## FIG. 21B